



Working Paper n. 11 - 2010

CORSI DI AZZERAMENTO DEFICIT DI COMPETENZE E RENDIMENTI DEGLI STUDENTI UNIVERSITARI

Maria De Paola

Dipartimento di Economia e Statistica
Università della Calabria
Ponte Pietro Bucci, Cubo 1/C
Tel.: +39 0984 492459
Fax: +39 0984 492421
e-mail: m.depaola@unical.it

Vincenzo Scoppa

Dipartimento di Economia e Statistica
Università della Calabria
Ponte Pietro Bucci, Cubo 1/C
Tel.: +39 0984 492464
Fax: +39 0984 492421
e-mail: v.scoppa@unical.it

Maggio 2010



Corsi di Azzeramento Deficit di Competenze e Rendimenti degli Studenti Universitari

Maria De Paola, Vincenzo Scoppa*

Dipartimento di Economia e Statistica, Università della Calabria

Abstract: In questo lavoro si analizza l'efficacia dei corsi a livello pre-universitari diretti a colmare eventuali lacune nelle competenze di base degli studenti. La nostra analisi riguarda gli effetti prodotti da un progetto che all'inizio dell'anno accademico 2008-09 ha offerto agli studenti immatricolati alle Università calabresi un insieme di attività formative. I tipici problemi di "self-selection" che si pongono nella valutazione degli effetti di interventi di questo tipo sono stati affrontati sia grazie alla disponibilità di dettagliate informazioni su caratteristiche, abilità e motivazioni degli studenti, sia cercando di identificare un gruppo di controllo costituito da studenti che non hanno partecipato all'intervento per fattori puramente casuali. Dalle nostre stime emerge un impatto positivo dei corsi di formazione sulle performance degli studenti nel primo anno di studi, sia usando modelli di regressione che stimatori di "matching". L'effetto positivo persiste, anche se si riduce quantitativamente, quando si considerano campioni di studenti più omogenei e gruppi di controllo costituiti da studenti che non hanno preso parte al progetto per fattori casuali.

*Keywords: Educational Production Function; Politiche formative; Istruzione Universitaria; matching estimators.
Classificazione JEL: I21; J31; D82*

1. Introduzione

La Regione Calabria, riconoscendo un ruolo centrale al capitale umano per lo sviluppo economico e sociale e la competitività del territorio, si è fatta promotrice all'inizio dell'anno accademico 2008-09 del progetto "Azzeramento Deficit di Competenze" che ha offerto agli studenti immatricolati alle Università calabresi un insieme di attività formative tese a colmare eventuali lacune nelle competenze di base.¹

Lo scopo di questo lavoro è quello di esaminare gli effetti prodotti da questo intervento sul rendimento accademico degli studenti iscritti all'Università della Calabria. Comprendere se interventi inseriti in uno stadio avanzato del processo formativo individuale possano portare ad un miglioramento nei risultati ottenuti dagli studenti è di cruciale importanza anche per

* E-mail: m.depaola@unical.it; v.scoppa@unical.it. Desideriamo ringraziare Alfio Cariola, Domenico Cersosimo, Federica Demaria, Alfredo Fortunato, Sara Laurita, Alessandra Perri e Michela Ponso per l'aiuto dato alla realizzazione della ricerca. Siamo grati all'Università della Calabria per averci consentito l'accesso ai dati amministrativi. Il lavoro ha beneficiato degli utili commenti di Giovanni Anania, Mariarosaria Agostino, Maria Teresa Traulo e Guglielmo Wolleb.

¹ Il progetto è stato finanziato dall'Assessorato Istruzione, Alta Formazione e Ricerca della Regione Calabria nell'ambito del "Piano Regionale per le Risorse Umane - Piano D'azione 2008" (POR Calabria, Fondo Sociale Europeo 2007-2013).

orientare altre politiche dell'istruzione, poiché molti studiosi ritengono che il sostegno alla formazione debba piuttosto concentrarsi sulle fasi iniziali del processo di istruzione ed essere diretto, quindi, alle scuole dell'infanzia e primarie (Carneiro e Heckman, 2003).

L'approccio che adottiamo consiste principalmente nel confrontare con gli opportuni strumenti statistici e econometrici le performance accademiche – misurate con il numero dei crediti maturati e il voto medio agli esami – degli studenti che hanno seguito le attività formative con le performance degli studenti che non hanno seguito affatto le attività formative o che ne hanno seguito solo una parte.

La principale difficoltà a cui si va incontro negli esercizi di valutazione di politiche di questo tipo è quella di separare l'impatto della politica da altri fattori "confondenti", quali le differenze pre-esistenti tra i soggetti beneficiari della politica ("trattati") e i soggetti che non sono stati interessati dall'intervento (i "non trattati"). Capire se le differenze che si registrano tra i destinatari dell'intervento e gli altri soggetti sono il risultato dell'intervento stesso o se sono dovuti, invece, ad altri elementi confondenti è in genere molto difficile, dato che non è possibile osservare contemporaneamente gli stessi soggetti nello status di beneficiari (nel gergo della letteratura, il cosiddetto "trattamento") e in quello di non-beneficiari (il "controfattuale"). Nel caso dei partecipanti al progetto "Azzeramento deficit di competenze", non è possibile osservare lo stesso studente nell'eventualità che abbia partecipato alle attività formative e nell'eventualità che non abbia partecipato.

Cercare di valutare la politica adottata confrontando i risultati ottenuti dal gruppo di studenti interessati dall'intervento con quelli del gruppo di studenti che non vi ha preso parte produce generalmente un problema di "distorsione da selezione" ("selection bias"): le due categorie di studenti potrebbero essere contraddistinte da particolari caratteristiche che non sono osservabili (al ricercatore) e che possono influenzare le performance accademiche. D'altra parte, problemi simili emergono quando si fa riferimento ai risultati ottenuti dallo stesso individuo prima e dopo l'intervento: tali differenze non forniscono tipicamente una stima attendibile degli effetti della politica, poiché nel corso del tempo possono cambiare anche altri fattori che incidono sulle variabili di interesse.

In linea di principio, per poter valutare l'impatto degli interventi di *policy* in modo rigoroso sarebbe necessario poter disporre di un gruppo di soggetti che viene sottoposto all'intervento ("gruppo di trattamento") e di un secondo gruppo di soggetti ("gruppo di controllo") che abbia caratteristiche del tutto analoghe al primo, ma che non viene sottoposto al "trattamento" in modo da poter essere utilizzato come metro di comparazione. Gli "esperimenti controllati randomizzati" assegnando, sulla base di una estrazione casuale, i soggetti al gruppo di trattamento e al gruppo di controllo assicurano la totale assenza di correlazione tra il trattamento e le caratteristiche osservabili e inosservabili degli individui e permettono di stimare gli effetti di una politica in modo non distorto (si veda Duflo et al. 2007) .

Tuttavia, in molte circostanze un gruppo di controllo di questo tipo non è disponibile e allora per poter valutare le politiche si ricorre a metodi non sperimentali che si basano sull'individuazione di un gruppo di confronto quanto più simile possibile al gruppo dei trattati. La stima dell'effetto dell'intervento si ottiene dalla differenza nei risultati di interesse tra il gruppo trattato e quello di confronto.

Nel presente esercizio di valutazione non disponiamo di un gruppo di controllo e abbiamo cercato di identificare l'effetto dell'intervento di policy utilizzando innanzitutto le numerose informazioni di cui disponiamo circa le caratteristiche e le abilità predeterminate degli studenti – usando sia metodi parametrici (modello di regressione OLS) che non parametrici (stimatori di “matching”). In secondo luogo, abbiamo cercato di costruire un gruppo di controllo che comprendesse solo quegli studenti che non hanno partecipato alle attività formative in esame per fattori casuali, esogeni, piuttosto che per scelta deliberata, in modo da avere un gruppo comparabile agli studenti “trattati”.

Dalla nostra analisi di regressione emerge che, controllando per un ampio numero di caratteristiche individuali, le attività formative previste dall'intervento “Azzeramento deficit di competenze” producono effetti ampiamente positivi e statisticamente significativi sui risultati accademici degli studenti. Risultati simili emergono anche utilizzando i cosiddetti stimatori di “matching” e “propensity score matching” – che confrontano i risultati degli studenti che hanno preso parte al progetto con quelli ottenuti da studenti con caratteristiche simili che però non hanno partecipato alle attività formative, considerando le differenze tra le medie (variamente ponderate) degli outcomes.

L'effetto delle attività formative rimane positivo anche quando, cercando di ovviare a problemi di selezione che potrebbero emergere se gli studenti che hanno partecipato all'intervento sono anche quelli con maggiore motivazione allo studio, si restringe l'analisi a gruppi di studenti più omogenei.

Un impatto positivo dell'intervento sulla performance degli studenti permane anche quando si considerano come gruppo di controllo solo quegli studenti che non hanno partecipato al progetto per fattori casuali. Infatti, sempre allo scopo di ovviare ai problemi di “*self-selection*”, seguendo l'approccio di Leuven e Oosterbeek (2008), per un sotto campione di studenti abbiamo condotto un sondaggio teso ad accertare quei casi in cui la mancata partecipazione al progetto è stata determinata da fattori casuali, non legati a fattori motivazionali.

Il nostro studio è collegato ad un'ampia letteratura che si è occupata della valutazione di programmi, finanziati con risorse pubbliche, tesi ad accrescere le competenze individuali (Ashenfelter, 1978; Ashenfelter e Card, 1978; LaLonde, 1986; Dehejia e Wahba, 1999). In questi esercizi di valutazione il problema della “distorsione da selezione” è cruciale, poiché gli individui si selezionano per il “trattamento”: in pratica, i corsi di formazione, previsti

nell'ambito dei diversi programmi, potrebbero essere seguiti dagli individui con le peggiori prospettive occupazionali (generando una distorsione verso il basso nelle stime dei rendimenti della formazione) oppure, in alternativa, potrebbero essere gli individui più motivati a intraprendere il percorso di formazione (generando una distorsione verso l'alto nelle stime).

LaLonde (1986) dà conto di un'ampia variabilità nei risultati ottenuti da diversi studi empirici e confrontando i risultati derivanti da dati sperimentali con quelli ottenuti da dati non sperimentali evidenzia le distorsioni presenti in questi ultimi. Dehejia e Wahba (1999) dimostrano che usando la tecnica del *propensity score matching* si riesce ad attenuare gli effetti di distorsione e grazie a questo metodo gli autori riescono a riprodurre risultati in linea con quelli prodotti dai dati sperimentali.

Alcune analisi tese a verificare l'efficacia dei corsi pre-universitari per compensare le lacune nelle competenze di base degli studenti sono state realizzate per gli Stati Uniti da Bettinger e Long (2009), Levin e Calcagno (2008) e da Calcagno e Long (2008) senza raggiungere risultati univoci.

Il nostro lavoro è anche collegato alla letteratura che ha cercato di analizzare gli effetti sui rendimenti degli studenti di una serie di input scolastici (dimensioni delle classi, abilità e esperienza degli insegnanti, ore di lezioni, caratteristiche dei "peer", risorse scolastiche, ecc.). Nonostante un considerevole volume di ricerche, i risultati ottenuti sono piuttosto ambigui e oggetto di accese dispute: mentre in una corposa rassegna Hanushek (1986) conclude che non sembra esistere una sistematica relazione tra input scolastici e performance degli studenti, Card e Krueger (1996) sono invece meno pessimisti².

L'articolo è organizzato nel seguente modo. Nel secondo paragrafo si descrivono le caratteristiche dell'intervento di *policy* e i dati utilizzati nelle analisi econometriche. Nel terzo paragrafo si presentano i risultati delle stime di diversi modelli di regressione, considerando come campione di riferimento tutti gli studenti iscritti all'Università della Calabria. Il quarto paragrafo presenta le stime dell'effetto del trattamento ottenute utilizzando il metodo degli stimatori di *matching* e *propensity score matching*. Nel paragrafo 5 presentiamo i risultati ottenuti quando si considerano campioni di studenti più omogenei in termini di motivazione. Nel sesto paragrafo si restringe il campione di riferimento in maniera da includere nel gruppo di controllo solo quegli studenti che hanno dichiarato di non aver partecipato al progetto per fattori casuali. Nel settimo paragrafo si offrono alcune brevi considerazioni conclusive.

2. Descrizione dell'intervento e dati

Le varie indagini internazionali dirette a misurare la qualità dell'istruzione e le competenze effettivamente acquisite dagli studenti residenti in diversi paesi e regioni (OCSE-PISA; IEA-

² Per una recente rassegna si veda Webbink (2005).

TIMSS; IEA-PIRLS; ecc.) mostrano per l'Italia livelli medi di conoscenze nettamente al di sotto della media dei paesi avanzati, in matematica, nelle scienze e nella comprensione verbale. Inoltre, dai risultati di queste indagini emerge inequivocabilmente un ampio divario di competenze tra gli studenti delle regioni centro-settentrionali e quelli delle regioni meridionali. Per esempio, dall'indagine PISA del 2006, posto pari a 500 il livello medio delle conoscenze per i paesi OCSE, la media italiana risulta pari a 475. Tuttavia, esiste una forte variabilità tra le diverse aree geografiche: mentre gli studenti del Nord-Est raggiungono una media di 520, la media per gli studenti calabresi (insieme a Basilicata, Sicilia e Sardegna) è di appena 432, un livello in linea con paesi in forte ritardo di sviluppo come Messico e Turchia. Come mostrano indagini empiriche che usano dati a livello individuale, i divari di competenze mostrati dagli studenti meridionali permangono anche controllando per una serie di variabili tradizionalmente considerate rilevanti nella spiegazione della performance scolastica, come le caratteristiche socio-economiche delle famiglie di provenienza (reddito, indicatori di ricchezza e istruzione dei genitori), le caratteristiche della scuola frequentata e le risorse investite.

Consapevole dei problemi derivanti da livelli inadeguati di capitale umano, la Regione Calabria ha finanziato, all'inizio dell'anno accademico 2008-09, un progetto denominato "Azzeramento Deficit di Competenze" che aveva l'obiettivo di rafforzare le competenze di base degli studenti immatricolati alle Università calabresi.³ Il progetto prevedeva una serie di attività formative della durata complessiva di 160 ore per il potenziamento di competenze di base a favore degli studenti iscritti al primo anno. Le attività formative, attuate mediante lezioni frontali e esercitazioni, hanno riguardato sia conoscenze di base in ambiti trasversali come la comprensione linguistica e la matematica, che discipline specifiche ad ogni singola Facoltà.

I corsi di potenziamento, incominciati all'inizio di Settembre 2008 (dopo che le competenze di base degli studenti sono state verificate attraverso un test iniziale) e conclusi alla fine di Ottobre 2008, sono stati tenuti sia da docenti universitari che da docenti esterni selezionati dalle Università. Agli studenti che hanno preso parte ad almeno l'80% delle attività formative e che hanno sostenuto un test finale, teso a valutare il livello di preparazione raggiunta, è stato corrisposto un voucher di 400 euro, a titolo di rimborso spese.

In questa analisi consideriamo solo gli studenti iscritti all'Università della Calabria, pari a circa 6.000, per i quali disponiamo di informazioni molto dettagliate su una serie di caratteristiche individuali (genere, età, voto di diploma, comune di residenza, tipologia di scuola frequentata, Facoltà scelta, ecc.), sui risultati ottenuti ai test somministrati all'inizio e alla fine dei corsi di azzeramento, sull'effettiva partecipazione alle attività formative (dati raccolti attraverso una procedura automatizzata appositamente predisposta per il monitoraggio delle

³ Un progetto con analoghe modalità è stato implementato anche nell'anno accademico 2009-10, ma è stato rivolto solo agli studenti che presentavano rilevanti deficit formativi.

presenze alle attività formative) e sulle carriere accademiche (esami sostenuti nel primo anno, crediti acquisiti, voti ottenuti agli esami).

Nella Tabella 1 sono riportate alcune statistiche descrittive relative a questo campione di studenti che comprende 6042 immatricolati iscritti a diverse Facoltà (il 23% a Lettere, il 22% ad Economia, il 15% ad Ingegneria, il 13,4% a Farmacia, il 12,3% a Scienze Matematiche Fisiche e Naturali e il 12% a Scienze Politiche). Circa il 59% degli studenti è rappresentato da donne. L'età media degli studenti è di circa 20 anni. Per quel che riguarda gli indicatori di abilità predeterminata, per tutti gli studenti appartenenti al campione osserviamo il voto di maturità e il tipo di scuola frequentata, mentre per un campione più ristretto di studenti osserviamo anche i risultati ottenuti ai test somministrati all'inizio dei corsi. Il voto medio al diploma è di 84,2, mentre la percentuale di risposte corrette alle sezioni del test relative rispettivamente alle competenze matematiche e alle competenze in italiano è del 38 e del 55 per cento. Il 36% degli studenti ha frequentato il Liceo Scientifico, il 12% il Liceo Classico, il 17% gli Istituti Tecnici Commerciali, il 7% il Liceo Socio-Psico-Pedagogico, mentre il rimanente 28% proviene da altri Istituti Tecnici e Professionali.

Il 71,5% ha frequentato almeno qualche ora dei corsi erogati nell'ambito del progetto in esame. In media gli studenti hanno seguito 80 delle 160 ore di formazione previste. Gli studenti che hanno seguito almeno l'80% delle lezioni, ottenendo così il voucher previsto dalla Regione Calabria, sono stati 3173, poco più del 50% dei potenziali beneficiari.

Tabella 1. Statistiche descrittive relative agli studenti immatricolati all'Università della Calabria nell'anno accademico 2008-09.

	Media	Deviazione standard	Min	Max	Osservazioni
Ore formazione	80.970	64.147	0	160	6042
Partecipazione80%	0.512	0.500	0	1	6042
Crediti	22.872	18.991	0	67	6042
Voto medio agli esami	24.083	3.017	18	31	4457
Donna	0.589	0.492	0	1	6042
Voto Maturità	84.265	12.185	60	100	6042
Scientifico	0.359	0.480	0	1	6042
Classico	0.132	0.338	0	1	6042
Istituto Commerciale	0.166	0.372	0	1	6042
Liceo Psico Pedagogico	0.074	0.261	0	1	6042
Altri Istituti Tecnici e Professionali	0.268	0.443	0	1	6042
Percentuale risposte corrette Matematica	0.384	0.188	0	1	4530
Percentuale risposte corrette Italiano	0.553	0.172	0	1	4721
Economia	0.221	0.415	0	1	6042
Farmacia	0.134	0.341	0	1	6042
Ingegneria	0.151	0.358	0	1	6042
Lettere	0.236	0.424	0	1	6042
Scienze Matematiche Fisiche e Naturali	0.123	0.329	0	1	6042
Scienze Politiche	0.118	0.323	0	1	6042
Età	20.494	4.200	17.717	62.327	6042
Anno di Maturità	2006.762	3.849	1969	2008	6042

I risultati degli studenti sono misurati in termini di crediti acquisiti nel corso del primo anno e di voti ottenuti agli esami (il voto di “30 e lode” è stato posto uguale a 31). I dati fanno riferimento agli esami sostenuti entro il 30 Settembre 2009, che rappresenta la conclusione del primo anno accademico.

Gli studenti iscritti al primo anno hanno acquisito in media 22.87 crediti, con un voto medio agli esami di 24.08 (questa media tiene conto solo degli esami superati). Da un’analisi preliminare risulta che il progetto è riuscito a coinvolgere soprattutto gli studenti con i risultati scolastici migliori. Infatti, le ore di frequenza alle attività formative sono crescenti all’aumentare del voto di maturità. Inoltre, risulta che gli studenti provenienti dai Licei Scientifici e Classici hanno seguito più assiduamente i corsi di formazione mentre bassa è stata la partecipazione degli studenti provenienti dagli Istituti Professionali. Si riscontra, inoltre, una scarsa partecipazione da parte degli studenti che in base ai test iniziali risultavano soffrire di importanti lacune nelle competenze linguistiche e matematiche.

Queste differenze nelle caratteristiche osservabili degli studenti “trattati” e “non trattati” segnalano che questi due gruppi di studenti potrebbero differire anche in termini di caratteristiche non osservabili e che, quindi, è necessario cercare di tenerne conto nel momento in cui si effettuano le stime.

In quanto segue presenteremo inizialmente i risultati ottenuti considerando come campione di riferimento tutti gli studenti immatricolati all’Università della Calabria e utilizzando lo stimatore dei minimi quadrati ordinari (OLS). Successivamente, la valutazione degli effetti dei corsi di azzeramento è compiuta utilizzando la tecnica dei *matching estimator*. In seguito per cercare di ovviare ai problemi di distorsione nelle stime dovute ai problemi di *self-selection* restringeremo la nostra analisi a campioni via via più omogenei, fino a considerare come gruppo di controllo solo quegli studenti che non sono riusciti a partecipare all’intervento di policy per motivi esogeni.

3. L’impatto dell’intervento sulla performance accademica degli studenti: le stime OLS

In questo paragrafo utilizzeremo diverse specificazioni di un modello di regressione standard per stimare l’impatto del progetto “Azzeramento Deficit di Competenze” sulla performance accademica degli studenti che vi hanno preso parte.

Il modello di regressione stimato è il seguente:

$$[1] \quad Y_i = \beta_0 + \beta_1 \text{Formazione}_i + \phi X_i + \psi_j + \varepsilon_i$$

La variabile dipendente, Y_i , rappresenta la performance accademica dello studente i ed è misurata, alternativamente, come numero di crediti formativi conseguiti durante il primo anno di studi universitari e come voto medio agli esami.

La variabile esplicativa di interesse, $Formazione_i$, nelle specificazioni principali rappresenta il numero totale di ore di attività formative seguite dallo studente i . Alternativamente, come indicatore di partecipazione alle attività formative si userà una variabile binaria (*Partecipazione80%*) che assume valore 1 per gli studenti che hanno seguito almeno l'80% delle attività formative previste nell'ambito del progetto (e zero altrimenti).⁴ X_i rappresenta un vettore di variabili individuali di controllo (misure di abilità pre-determinate e altre caratteristiche individuali); ψ_j include le dummy di Facoltà dello studente, per tener conto di eventuali eterogeneità nella difficoltà dei diversi percorsi di studio o nei metri di valutazione adottati; ε_i è un termine di errore che cattura shock idiosincratici o caratteristiche inosservabili dello studente.

Nella Tabella 2 la variabile dipendente è rappresentata dai crediti maturati, mentre le variabili esplicative includono, oltre alle ore di formazione seguite (colonne 1-3) e la dummy *Partecipazione80%* (colonne 3-6), alcune variabili individuali quali il genere, l'età, il tipo di scuola frequentata, il voto di maturità, l'anno di conseguimento della maturità e la provincia di residenza. Le stime sono state effettuate con il metodo dei minimi quadrati (OLS). Nella Tabella sono riportati i coefficienti stimati e i relativi *standard error* (tra parentesi). La presenza di eteroschedasticità negli errori è corretta con la procedura di White-Huber: gli standard error sono così robusti all'eteroschedasticità.

Come si può notare dalla colonna 1 della Tabella 2, le ore di formazione complessivamente seguite dagli studenti hanno avuto un impatto fortemente positivo sul conseguimento di crediti durante il primo anno di Università: per un generico studente, aver seguito 100 ore di formazione in più determina un incremento del numero medio di crediti conseguiti nel primo anno pari a 6,7. Tale effetto è confortato da un'alta significatività statistica (la statistica t è pari a 19). Considerando che la media dei crediti è stata pari a 22,9, in termini percentuali l'effetto di 100 ore di attività formative corrisponde ad un aumento nel numero di crediti maturati del 29%.

Si noti che la stima di tale effetto è ottenuta tenendo costanti un'ampia serie di misure di abilità e caratteristiche individuali osservabili e quindi non è distorta per via della correlazione esistente, ad esempio, tra voto di diploma e intensità nella frequenza ai corsi di formazione. Tuttavia, lo stimatore di β_1 potrebbe essere distorto a causa dell'esistenza di altre

⁴ Seguire l'80% delle lezioni costituiva per gli studenti un requisito necessario per l'ottenimento del voucher di 400 euro.

caratteristiche inosservabili dello studente, come la motivazione, che potrebbero essere sia correlate alla partecipazione ai corsi di formazione che influenzare i risultati accademici.

Per esaminare l'impatto dell'intervento sulla performance di studenti caratterizzati da livelli di abilità differenti, nelle colonne 2 e 3 si presentano i risultati delle stime OLS quando si suddivide il campione in due gruppi usando quale variabile discriminante il voto di maturità. In particolare sono stati denominati "studenti ad alte abilità", coloro i quali hanno conseguito il diploma di scuola secondaria con un voto superiore al voto medio di 84, e "studenti a basse abilità" coloro i quali hanno ottenuto un voto di maturità inferiore alla media. Rientrano nel primo gruppo 3037 studenti e nel secondo 3005. I risultati delle stime mostrano che l'effetto della formazione è positivo e significativo sia per gli studenti ad alte abilità (7.6 crediti in più per 100 ore di formazione) sia per gli studenti a basse abilità (+5.8 crediti), ma risulta maggiore per gli studenti con voto di maturità più alto.

Nelle colonne 4-6 della Tabella 2 si misura la partecipazione degli studenti all'intervento formativo – piuttosto che con il numero di ore di attività seguite – usando la dummy *Partecipazione*80%, cioè considerando come "trattati" gli studenti che hanno partecipato ad almeno l'80% delle attività formative (che corrisponde a 128 ore). E' possibile, infatti, che la relazione tra ore di formazione e performance accademica non sia lineare e che sia necessario raggiungere una certa soglia per realizzare miglioramenti significativi nei rendimenti.

I risultati delle nuove regressioni sono abbastanza simili qualitativamente alle specificazioni delle colonne 1-3. Come si può notare dalla colonna 4 della Tabella 2, gli studenti che hanno seguito almeno 128 ore di formazione registrano un incremento nel numero di crediti conseguiti durante il primo anno pari a 8.7. Nelle colonne 5 e 6 trova conferma il risultato emerso considerando come variabile dipendente il numero di ore di formazione seguite: l'effetto dell'intervento di policy è maggiore sugli studenti con voto di maturità superiore alla media (+9.7 crediti), sebbene sia considerevole anche per gli studenti con minori abilità (+7.6 crediti).

Per quel che riguarda l'impatto esercitato dalle caratteristiche individuali si riscontrano risultati che confermano quanto già emerso nella letteratura di economia dell'istruzione (si veda, per esempio, Stinebrickner e Stinebrickner, 2008; Fuchs e Woessmann, 2007). A parità di altre caratteristiche, il voto di maturità ha un impatto fortemente positivo, e statisticamente significativo, sul conseguimento di crediti: studenti con un voto di maturità maggiore di 10 punti conseguono circa 5 crediti in più durante il primo anno di Università (+22%). Gli studenti che hanno frequentato il Liceo Scientifico o Classico maturano un numero di crediti maggiore rispetto a coloro che hanno frequentato una scuola professionale (che rappresenta la categoria di riferimento). Per quanto riguarda il genere, emerge che le donne hanno un rendimento migliore. L'età risulta essere una variabile non significativa per spiegare il fenomeno che stiamo analizzando, mentre l'aver conseguito la maturità nell'anno di immatricolazione all'Università è

associato a una migliore performance. Infine, vi sono differenze rilevanti a seconda della Facoltà a cui lo studente è iscritto: gli studenti iscritti alla Facoltà di Economia (categoria di riferimento) maturano un maggior numero di crediti rispetto agli studenti iscritti a tutte le altre Facoltà, la differenza è statisticamente significativa fatta eccezione per gli studenti iscritti alla Facoltà di Scienze Politiche.

Tabella 2. Relazione tra crediti conseguiti e ore di formazione: i risultati delle stime OLS

	(1)	(2)	(3)	(4)	(5)	(6)
	Tutti	Studenti ad alte abilità	Studenti a basse abilità	Tutti	Studenti ad alte abilità	Studenti a basse abilità
Ore formazione	0.067*** (0.004)	0.076*** (0.006)	0.058*** (0.004)			
Partecipazione80%				8.712*** (0.451)	9.713*** (0.704)	7.616*** (0.568)
Donna	1.934*** (0.479)	1.218* (0.740)	2.114*** (0.606)	1.964*** (0.477)	1.299* (0.736)	2.109*** (0.604)
Voto Maturità	0.510*** (0.019)	0.707*** (0.061)	0.429*** (0.038)	0.509*** (0.019)	0.718*** (0.060)	0.425*** (0.038)
Scientifico	7.125*** (0.541)	7.750*** (0.815)	6.616*** (0.698)	7.089*** (0.541)	7.640*** (0.815)	6.615*** (0.696)
Classico	3.245*** (0.728)	3.714*** (1.033)	2.857*** (0.990)	3.205*** (0.727)	3.687*** (1.032)	2.780*** (0.991)
Istituto Tecnico Commerciale	1.235* (0.642)	0.177 (1.057)	1.808** (0.776)	1.096* (0.642)	0.168 (1.057)	1.612** (0.777)
Liceo Socio-Psico-Ped	1.117 (0.912)	0.154 (1.395)	1.726 (1.156)	1.116 (0.909)	0.103 (1.383)	1.737 (1.158)
Farmacia	-7.073*** (0.662)	-10.366*** (1.014)	-4.309*** (0.843)	-6.792*** (0.664)	-9.951*** (1.019)	-4.117*** (0.847)
Ingegneria	-6.381*** (0.741)	-6.397*** (0.994)	-8.894*** (1.103)	-5.920*** (0.734)	-5.986*** (0.986)	-8.231*** (1.084)
Lettere	-2.677*** (0.672)	-3.162*** (1.010)	-2.003** (0.882)	-2.300*** (0.671)	-2.839*** (1.005)	-1.582* (0.885)
Scienze Matematiche Fisiche e Naturali	-3.528*** (0.740)	-4.315*** (1.127)	-3.386*** (0.939)	-3.493*** (0.740)	-3.991*** (1.124)	-3.572*** (0.938)
Scienze Politiche	-0.244 (0.784)	2.954** (1.389)	-1.502 (0.959)	-0.132 (0.779)	3.341** (1.379)	-1.465 (0.955)
Anno diploma	0.476*** (0.184)	0.989*** (0.266)	0.276 (0.240)	0.543*** (0.186)	1.094*** (0.281)	0.313 (0.234)
Età	0.062 (0.169)	-0.002 (0.238)	0.100 (0.225)	0.079 (0.171)	0.036 (0.254)	0.093 (0.218)
Costante	-980.793*** (372.999)	-2,026.169*** (537.649)	-574.451 (486.879)	-1,114.197*** (376.749)	-2,238.248*** (568.130)	-647.826 (474.879)
Osservazioni	6042	3037	3005	6042	3037	3005
R-quadro	0.276	0.220	0.175	0.279	0.223	0.176

Note: Stime OLS sul totale degli immatricolati. Variabile dipendente: Crediti maturati dagli studenti. In tutte le regressioni controlliamo per le dummy per le province di residenza degli studenti. Gli *standard error* (riportati tra parentesi) sono robusti all'eteroschedasticità. I simboli ***, **, * indicano, rispettivamente, significatività al 1%, 5% e 10%.

La performance degli studenti può essere utilmente misurata anche dal voto medio che essi hanno conseguito agli esami sostenuti durante il primo anno di studi (il voto medio è calcolato considerando solo gli esami superati e ponderando per il numero di crediti

corrispondente a ciascun esame). Nella Tabella 3 sono riportati i risultati delle stime OLS di alcuni modelli che analizzano la relazione tra la frequenza ai corsi di azzeramento e il voto medio ottenuto. Nelle prime tre colonne la partecipazione all'intervento è misurata considerando il numero di ore di attività formative a cui lo studente ha preso parte, mentre nelle colonne 4, 5 e 6 si considera la dummy *Partecipazione80%*.

Tabella 3. Relazione tra voto medio conseguito agli esami e ore di formazione: i risultati delle stime OLS

	(1) Tutti	(2) Studenti ad alte abilità	(3) Studenti a basse abilità	(4) Tutti	(5) Studenti ad alte abilità	(6) Studenti a basse abilità
Ore formazione	0.002*** (0.001)	0.002** (0.001)	0.002** (0.001)			
Partecipazione80%				0.231*** (0.078)	0.158 (0.106)	0.315*** (0.114)
Donna	0.250*** (0.085)	0.128 (0.115)	0.379*** (0.124)	0.250*** (0.085)	0.130 (0.116)	0.381*** (0.124)
Voto Maturità	0.098*** (0.003)	0.145*** (0.009)	0.072*** (0.008)	0.098*** (0.003)	0.146*** (0.009)	0.072*** (0.008)
Scientifico	1.249*** (0.094)	1.530*** (0.124)	0.915*** (0.143)	1.250*** (0.094)	1.534*** (0.124)	0.911*** (0.143)
Classico	1.196*** (0.120)	1.335*** (0.147)	1.030*** (0.206)	1.198*** (0.120)	1.345*** (0.147)	1.021*** (0.206)
Istituto Tecnico Commerciale	0.223* (0.120)	0.217 (0.169)	0.215 (0.171)	0.222* (0.120)	0.227 (0.169)	0.205 (0.171)
Liceo Socio-Psico-Ped	0.394*** (0.143)	0.320* (0.181)	0.441* (0.225)	0.394*** (0.143)	0.326* (0.181)	0.436* (0.225)
Farmacia	-0.063 (0.119)	0.410** (0.164)	-0.631*** (0.172)	-0.057 (0.119)	0.421** (0.164)	-0.622*** (0.172)
Ingegneria	0.515*** (0.128)	0.503*** (0.155)	0.242 (0.235)	0.531*** (0.127)	0.529*** (0.154)	0.240 (0.235)
Lettere	3.449*** (0.110)	3.449*** (0.140)	3.418*** (0.172)	3.464*** (0.109)	3.472*** (0.139)	3.430*** (0.171)
Scienze Matematiche Fisiche e Naturali	0.660*** (0.128)	0.671*** (0.160)	0.466** (0.201)	0.660*** (0.128)	0.682*** (0.161)	0.464** (0.200)
Scienze Politiche	1.433*** (0.141)	1.886*** (0.214)	1.054*** (0.191)	1.439*** (0.141)	1.904*** (0.214)	1.050*** (0.191)
Anno diploma	-0.080** (0.033)	0.025 (0.044)	-0.162*** (0.031)	-0.077** (0.034)	0.032 (0.045)	-0.163*** (0.032)
Età	0.022 (0.030)	0.087** (0.037)	-0.054** (0.026)	0.022 (0.031)	0.087** (0.037)	-0.056** (0.026)
Costante	174.031*** (67.427)	-43.191 (89.770)	342.911*** (63.410)	168.837** (67.855)	-57.123 (89.985)	344.380*** (64.217)
Osservazioni	4457	2466	1991	4457	2466	1991
Pseudo R-quadro	0.389	0.328	0.376	0.388	0.327	0.376

Note: Stime OLS sul totale degli immatricolati. Variabile dipendente: Voto medio agli esami superati. In tutte le regressioni controlliamo per le dummy per le province di residenza degli studenti. Gli *standard error* (riportati tra parentesi) sono robusti all'eteroschedasticità. I simboli ***, **, * indicano, rispettivamente, significatività al 1%, 5% e 10%.

Dalle stime emerge che le ore di formazione hanno avuto un impatto positivo e statisticamente significativo sul voto medio ottenuto dagli studenti agli esami superati. Per

esempio, gli studenti che hanno seguito 100 ore di formazione ottengono un voto medio più alto di 0.2 punti. Non si riscontrano differenze significative nella grandezza dell'effetto per studenti caratterizzati da differenti livelli di abilità.

Gli studenti che hanno partecipato ad almeno l'80% delle attività formative, ottengono un voto medio agli esami di circa 0.23 punti superiore a quello ottenuto dagli studenti che non hanno frequentato i corsi o che hanno frequentato un numero di ore inferiore a 128.⁵

4. La valutazione dell'impatto dei corsi con tecniche di “matching” e “propensity score matching”

In questo paragrafo, allo scopo di attenuare il problema della distorsione da selezione (“selection bias”) e cercare di individuare un contraffattale appropriato per il gruppo di studenti trattati, useremo strategie di stima alternative al modello di regressione: le tecniche dei “matching” e del “propensity score matching”, sviluppate, tra gli altri, da Rosenbaum e Rubin (1983), Angrist (1998) e Abadie e Imbens (2006) e diffusamente adottate di recente negli esercizi di valutazione econometrica.

L'idea alla base di queste tecniche è quella di comparare individui “trattati” solo con individui “non trattati” che hanno caratteristiche osservabili molto simili. La stima dell'effetto medio del trattamento (*Average Treatment Effect* o *ATE*) o dell'effetto medio del trattamento sui trattati (*Average Treatment Effect on Treated* o *ATT*) è ottenuta ponderando opportunamente le differenze tra le medie dell'outcome dei soggetti trattati rispetto all'outcome dei soggetti non trattati, all'interno di ogni gruppo di soggetti che hanno in comune gli stessi valori (o valori molto simili) delle covariate.

Per esempio, nel caso che stiamo analizzando gli studenti possono essere “abbinati” (*matched*) sulla base del voto di diploma, del tipo di scuola frequentata, del genere, della Facoltà a cui sono iscritti, ecc.. All'interno di ogni combinazione di caratteristiche (celle)⁶ si calcola la differenza tra la media dei crediti acquisiti dagli studenti trattati e la media dei crediti ottenuti dagli studenti non trattati. L'effetto del trattamento sui trattati si ottiene ponderando le differenze in ogni cella per il corrispondente numero dei trattati.

Gli stimatori di *matching* identificano l'effetto causale del trattamento e sono consistenti sotto l'ipotesi di “Indipendenza Condizionale” (“*Conditional Independence*”

⁵ Risultati molto simili si ottengono stimando un modello Tobit per tener conto del fatto che il voto medio agli esami è osservato solo per gli studenti che hanno superato almeno un esame e che, quindi, hanno ottenuto un voto almeno pari a 18, mentre per tutti gli altri studenti, cioè per quelli che non hanno superato esami, non risulta alcun voto.

⁶ Ad esempio, una cella è costituita dagli studenti con voto di maturità pari a 100, provenienti dal Liceo Scientifico, di sesso maschile, iscritti alla Facoltà di Economia.

Assumption”), cioè sotto l’ipotesi che, condizionatamente alle caratteristiche osservabili degli studenti, il trattamento è come se fosse assegnato in modo casuale.⁷

Rispetto al modello di regressione, il metodo dei *matching* ha due vantaggi: 1) non è necessario specificare la forma funzionale della relazione che lega la variabile dipendente alle variabili esplicative, quindi il metodo dei *matching* (non parametrico) non poggia sull’ipotesi di linearità come il modello di regressione; 2) nella determinazione dell’effetto del trattamento, i *matching* utilizzano solo le osservazioni per le quali le unità trattate sono simili alle unità non trattate, mentre ignorano le osservazioni per le quali i soggetti trattati sono molto diversi da quelli non trattati, cioè ignorano le osservazioni che non hanno un “supporto comune”.

I metodi di matching più diffusi consentono di calcolare l’effetto del trattamento quando il trattamento stesso è rappresentato da una variabile dicotomica (partecipazione o non partecipazione all’intervento), mentre le tecniche per considerare casi in cui il trattamento è rappresentato da una variabile continua sono ancora poco sviluppate e correntemente oggetto di ricerca.

Di conseguenza, verificheremo l’impatto prodotto dalla partecipazione ad almeno l’80% delle lezioni previste dall’intervento di policy sui crediti acquisiti dagli studenti e sul voto medio agli esami.

Per stimare l’effetto medio del trattamento sui trattati (*ATT*) sui crediti e sul voto medio usiamo lo specifico stimatore di *matching* sviluppato in Abadie e Imbens (2006) e implementato per Stata da Abadie *et al.* (2004) che si basa sulla tecnica del “Nearest Neighbor” che accoppia ad ogni unità trattata una (o più) unità non trattata “più vicina” sulla base di una funzione di distanza che considera le seguenti caratteristiche degli studenti: il voto di diploma, il genere, l’età, la scuola di provenienza, l’anno di maturità, la provincia di residenza e la Facoltà. La procedura imputa l’outcome potenziale non realizzato (il controfattuale) ad ogni studente trattato usando l’outcome ottenuto dallo studente appartenente al gruppo di controllo che sulla base delle variabili di matching presenta la distanza minore dallo studente trattato. Nel caso di matching multipli, diciamo m , si usa come outcome potenziale la media degli outcome degli m studenti più vicini.

Formalmente, l’*ATT* può essere espresso come:

$$\tau = \frac{1}{n_T} \sum_{i=1}^{n_T} (Y_i^T - Y_{(i)}^C)$$

⁷ Bisogna però evidenziare che gli stimatori basati sui matching non permettono di ovviare a problemi di distorsione nelle stime che derivano dalla selezione degli individui nel trattamento in relazione a caratteristiche non osservabili.

dove Y_i^T rappresenta l'outcome dell' i -esimo studente trattato, mentre $Y_{(i)}^C$ rappresenta l'outcome dello studente nel gruppo di controllo associato allo studente i sulla base della similarità delle covariate, n_T è il numero totale degli studenti trattati.

Il metodo permette di usare le osservazioni per il *matching* più di una volta, cosicché occorre decidere quanti individui non trattati abbinare ad ogni individuo trattato. La scelta presenta un trade-off: aumentando il numero di matches è possibile ridurre la varianza, ma poiché in questo modo si tende ad abbinare individui via via meno “simili” si tende anche ad aumentare il grado di distorsione. Per verificare se i risultati sono robusti rispetto alla scelta del numero dei matches, abbiamo usato tre diverse procedure: un singolo match, 4 e 8 matches. Poiché, come mostrano Abadie e Imbens (2006), il semplice stimatore di matching è distorto in campioni finiti quando il matching non è perfetto, stimiamo l'ATT utilizzando l'opzione di “bias-adjustment” che corregge questa distorsione. Inoltre, gli standard errors sono robusti all'eteroschedasticità.

La Tabella 4 riporta le stime dell'effetto medio del trattamento sui trattati. Nella prima colonna riportiamo le stime per quanto riguarda i crediti e nella seconda colonna valutiamo l'effetto sul voto medio.

L'effetto medio della partecipazione ad almeno l'80% delle attività formative previste dal progetto sul numero di crediti conseguiti è positivo – variando da 7.68 a 8.37 a seconda del numero dei matches – e risulta solo lievemente inferiore all'effetto riscontrato attraverso lo stimatore OLS (8.71, Tabella 2, colonna 4).⁸ L'effetto ha un livello molto alto di significatività statistica ($t=12.3$).

Un impatto positivo delle attività formative si riscontra anche sul voto medio agli esami (da 0.16 a 0.27), effetto che è in linea, anche in questo caso, con la stima OLS.

Tabella 4. Stime dell'effetto medio del trattamento sui trattati (“Average Treatment Effect on the Treated”) della partecipazione ai corsi sui crediti conseguiti e sul voto medio agli esami

	Crediti	Voto medio
(1 match)	8.198*** (0.665)	0.167* (0.101)
(4 matches)	8.375*** (0.563)	0.263*** (0.092)
(8 matches)	7.680*** (0.530)	0.269*** (0.089)
Sample	Tutti gli immatricolati	Immatricolati che hanno superato almeno un esame
Osservazioni	6042	4457

Note: La tabella riporta le stime dell'effetto medio del Trattamento sui Trattati (ATT) relativo alla partecipazione all'80% delle lezioni (*Partecipazione80%*). Variabili usate per il matching: Voto maturità; Donna; Liceo Scientifico; Liceo Classico; Istituto Tecnico Commerciale; Liceo Psico-pedagogico; Dummy per province di residenza; Anno maturità; Dummy di Facoltà. Il simbolo *** indica significatività all'1%

⁸ Abbiamo anche stimato l'effetto medio del trattamento (ATE) e l'effetto del trattamento per i non trattati (ATC) (non riportati) trovando risultati sostanzialmente simili all'ATT.

Risultati simili a quelli presentati nella Tabella 4 (che non riportiamo) si riscontrano anche quando si abbinano soggetti trattati e non trattati utilizzando le tecniche basate sul *propensity score matching*, cioè confrontando gli outcome di soggetti trattati e non trattati abbinati sulla base della loro probabilità di ottenere il trattamento (il cosiddetto *score*).

5. Alcuni esercizi di robustezza: le stime su campioni di studenti più omogenei

Nonostante nell'analisi svolta nel precedente paragrafo si sia tenuto conto di un gran numero di caratteristiche osservabili degli studenti che ne influenzano i risultati accademici e si sia cercato attraverso le tecniche dei *matching* di confrontare soggetti con caratteristiche simili, le stime potrebbero ancora soffrire di distorsioni indotte da un processo di selezione nella partecipazione al progetto basato su fattori inosservabili che influenzano anche la performance accademica. Sebbene individui simili in termini di caratteristiche osservabili quali genere, tipo di scuola frequentata, abilità tendono tipicamente ad essere simili anche in termini di caratteristiche inosservabili, non è possibile escludere che a parità di caratteristiche, la partecipazione all'intervento di policy sia legata a qualche differenza nelle caratteristiche non osservabili. Nel caso in esame potrebbe accadere, ad esempio, che a parità di caratteristiche osservabili solo gli studenti più motivati partecipino al progetto e, di conseguenza, l'effetto stimato sulla performance potrebbe essere, almeno in parte, il risultato della maggiore motivazione degli studenti piuttosto che essere determinato esclusivamente dalle attività formative previste dall'intervento di policy.

Allo scopo di attenuare il problema della distorsione da selezione che potrebbe condurre a una sovrastima dell'effetto del progetto "Azzeramento deficit di competenze", nella Tabella 5 riportiamo alcune regressioni effettuate su campioni di studenti con caratteristiche più omogenee, che ci portano a ritenere che siano più simili anche in termini di motivazione allo studio.

In particolare, per cercare di includere nel campione di riferimento solo soggetti con livelli di motivazione simili abbiamo considerato due sotto-campioni, uno composto solo dagli studenti che hanno preso parte al test iniziale che è stato somministrato all'inizio dell'anno accademico (*Campione Test*) (il campione si riduce da 6042 a 4435) e l'altro che include gli studenti che hanno partecipato al test iniziale e che hanno anche risposto a un questionario di indagine sugli immatricolati che veniva proposto on line al momento della domanda di immatricolazione (*Campione Test+Questionario*) (campione di 3435 studenti).

Nelle colonne 1 e 2 della Tabella 5 si riportano le stime relative all'effetto rispettivamente delle ore di formazione e della dummy *Partecipazione80%* sul campione degli

studenti che ha preso parte al test iniziale. Questo campione di studenti è più omogeneo in termini di motivazione, poiché sono esclusi gli studenti che – con la loro mancata partecipazione al test (che serviva a rilevare eventuali lacune nelle competenze di base) – hanno segnalato uno scarso interesse verso il percorso di studi intrapreso. Un ulteriore vantaggio derivante dall'utilizzo di questo campione consiste nel fatto che i risultati ottenuti al test iniziale ci permettono di controllare per due ulteriori misure di abilità predeterminate degli studenti, cioè la percentuale di risposte esatte alle domande della sezione del test riguardante le competenze in matematica e la percentuale di risposte esatte relative alla conoscenza della lingua italiana.

Come si può notare nelle colonne 1 e 2 della Tabella 5, in cui si riportano le stime OLS usando come variabile dipendente il numero di crediti acquisiti dagli studenti, anche considerando questo campione di studenti più omogeneo l'effetto delle ore di formazione rimane positivo e statisticamente significativo (all'1%, con una statistica $t=13.9$). Per questo campione di studenti l'impatto di 100 ore di formazione risulta pari a 6.3 crediti (colonna 1) piuttosto che 6.7 (la stima ottenuta considerando tutti gli immatricolati, Tabella 2, colonna 1). La partecipazione all'80% delle lezioni consente di incrementare di quasi 8 i crediti acquisiti (colonna 2), effetto di poco inferiore rispetto alle stime ottenute usando il campione con tutti gli immatricolati (8.7).

Nelle colonne 3 e 4 si considera il campione di studenti che non solo ha partecipato al test iniziale, ma ha anche risposto al questionario di indagine sugli immatricolati. Restringere l'analisi a questo campione consente di utilizzare alcune informazioni sul background familiare dello studente (ottenute attraverso l'Indagine sugli Immatricolati), in particolare è possibile controllare per il livello di istruzione dei genitori. Il coefficiente che misura l'impatto delle ore di formazione (colonna 3) diventa in questo caso pari a 0.054 (significativo all'1%): gli studenti che hanno seguito 100 ore delle attività formative previste dal progetto conseguono 5.4 crediti in più durante il primo anno di studi rispetto agli studenti che non hanno seguito alcuna lezione. Nella colonna 4 usiamo come variabile esplicativa la dummy *Partecipazione80%*: i risultati mostrano che aver seguito almeno l'80% delle attività formative previste dall'intervento produce un effetto positivo sui crediti conseguiti di 6.83 (statisticamente significativo all'1%). Pertanto, quando si considera questo campione più omogeneo di studenti si continua a stimare un effetto positivo dell'intervento anche se esso risulta minore rispetto alle stime precedenti.

Un'analisi analoga è stata condotta considerando come variabile dipendente il voto medio agli esami sostenuti (queste stime non sono riportate e sono disponibili su richiesta). L'effetto delle attività formative continua ad essere statisticamente significativo all'1%, ma è quantitativamente minore rispetto all'effetto individuato considerando come campione di riferimento tutti gli studenti immatricolati (vedi Tabella 3).

Tabella 5. L'impatto delle ore di formazione sui crediti formativi. Regressioni OLS su sample più omogenei di studenti.

	(1) Campione Test	(2) Campione Test	(3) Campione Test+Questionario	(4) Campione Test+Questionario
Ore formazione	0.063*** (0.005)		0.054*** (0.006)	
Partecipazione80%		7.919*** (0.548)		6.832*** (0.665)
Donna	2.661*** (0.563)	2.591*** (0.559)	1.881*** (0.660)	1.907*** (0.656)
Voto Maturità	0.444*** (0.024)	0.445*** (0.024)	0.482*** (0.027)	0.482*** (0.027)
Scientifico	4.892*** (0.675)	4.845*** (0.674)	5.463*** (0.796)	5.424*** (0.796)
Classico	2.074** (0.884)	1.971** (0.881)	2.882*** (1.028)	2.772*** (1.028)
Istituto Tecnico Comm.	0.804 (0.756)	0.671 (0.755)	0.887 (0.883)	0.834 (0.882)
Liceo Socio-Psico-Ped	1.499 (1.117)	1.484 (1.115)	1.057 (1.319)	0.937 (1.316)
Percentuale Risposte Corrette Matematica	16.058*** (1.629)	15.688*** (1.630)	16.137*** (1.832)	15.758*** (1.832)
Percentuale Risposte Corrette Italiano	5.195*** (1.719)	5.418*** (1.710)	3.948** (1.979)	4.279** (1.968)
Istruzione Padre			-0.030 (0.088)	-0.001 (0.088)
Istruzione Madre			-0.112 (0.088)	-0.131 (0.087)
Costante	-1,454.746*** (410.592)	-1,584.141*** (426.210)	-1,567.880*** (512.412)	-1,680.527*** (525.378)
Osservazioni	4435	4435	3435	3435
R-quadro	0.294	0.298	0.279	0.282

Note: Stime OLS su campioni più omogenei di studenti. In tutte le regressioni controlliamo per dummy di Facoltà, dummy di provincia di residenza, età e anno di conseguimento della maturità. Gli *standard error* sono robusti all'eteroschedasticità. I simboli ***, **, * indicano, rispettivamente, significatività al 1%, 5% e 10%.

6. Le stime basate sul confronto tra trattati e un gruppo di controllo casuale

In questo paragrafo, per evitare i problemi di selezione che si presentano se gli studenti che hanno partecipato all'intervento di policy sono anche quelli con caratteristiche inosservabili che incidono positivamente sulla performance accademica, abbiamo adottato la strategia proposta da Leuven e Oosterbeeck (2008) basata sull'idea di restringere il gruppo di controllo in modo da includere solo quei soggetti che avevano intenzione di partecipare all'intervento di policy, ma che non lo hanno fatto a causa di imprevisti che li hanno costretti a rinunciare. L'idea alla base della strategia di identificazione è che questi soggetti non differiscono, in termini di motivazione e altre caratteristiche inosservabili, dagli studenti effettivamente trattati.

A tale scopo su un sottocampione di 462 studenti (quelli iscritti nell'anno accademico 2008-09 al Corso di Laurea in Economia Aziendale offerto dalla Facoltà di Economia dell'Università della Calabria), al termine dell'anno accademico, abbiamo effettuato una

indagine aggiuntiva attraverso un questionario che chiedeva agli studenti di fornire una serie di informazioni relative al loro coinvolgimento nel progetto “Azzeramento deficit di competenze” e in caso di mancata partecipazione alle attività formative chiedeva di specificarne i motivi.

Per rispondere alla domanda relativa alla mancata partecipazione al progetto gli studenti potevano scegliere tra le seguenti risposte: 1) Non ho partecipato perché ci sono stati degli imprevisti che me lo hanno impedito; 2) Non ho partecipato perché non avevo tempo; 3) Non ho partecipato perché il voucher era troppo basso; 4) Non ho partecipato perché non lo ritenevo utile; 5) Non ho partecipato perché, nonostante abbia cercato, non ho trovato un alloggio vicino all'Università.

Gli studenti che hanno risposto al questionario sono stati in totale 364. Di questi studenti 277 hanno partecipato al progetto “Azzeramento deficit di competenze”, mentre i rimanenti 87 non vi hanno preso parte. Degli 87 che non vi hanno preso parte, 64 dichiarano di non aver partecipato a causa di imprevisti o perché non sono riusciti a trovare un alloggio vicino all'Università. Di conseguenza, abbiamo creato un gruppo di controllo in cui sono inclusi solo i 64 studenti che dichiarano di non aver preso parte al progetto per fattori casuali e per i quali, quindi, la mancata partecipazione non può essere imputata a una scelta deliberata.

Nella Tabella 6 vengono presentati i risultati dell'impatto prodotto dal progetto in esame considerando diversi campioni di studenti. La variabile dipendente è costituita dalla dummy *Partecipazione80%* che prende valore uno per gli studenti che hanno preso parte ad almeno l'80% delle attività formative previste dal progetto. Nella colonna (1) si riportano i risultati che si ottengono quando si considera tutto il campione degli studenti iscritti al Corso di Laurea in Economia Aziendale. L'effetto dell'intervento su questi studenti è in linea con quello stimato quando sono stati considerati tutti gli studenti immatricolati all'Università della Calabria (8.47 invece di 8.71, si veda Tabella 2). Nelle colonne 2 e 3 si considerano rispettivamente il campione degli studenti che ha partecipato al test iniziale (*Campione Test*) e quello che oltre ad aver partecipato al test iniziale ha risposto al questionario immatricolati (*Campione Test+Questionario*). L'effetto stimato nei due casi è pari rispettivamente a 7.85 e 6.82, ancora del tutto in linea con l'effetto stimato per tutti gli studenti immatricolati all'Università della Calabria (si veda Tabella 5).

Nella colonna 4, invece, si considerano come gruppo di controllo esclusivamente quegli studenti che dichiarano di non aver partecipato al progetto per motivi casuali. Pertanto il campione è costituito da tutti gli studenti trattati più i 64 studenti che hanno dichiarato di non essere riusciti a seguire le attività formative. La stima dell'effetto dell'intervento si riduce notevolmente: gli studenti che hanno partecipato all'80% delle attività formative conseguono in media circa 3.2 crediti in più rispetto agli studenti che non hanno partecipato affatto al progetto o che hanno frequentato un numero di ore di attività formative inferiore alla soglia dell'80%. La significatività statistica si riduce – anche a causa del ridotto numero di osservazioni – al 10%.

Questi risultati suggeriscono che le attività formative del progetto hanno avuto un impatto positivo, poiché l'effetto continua a rimanere positivo nonostante si confrontino studenti via via più omogenei. Tuttavia, è probabile che l'effetto stimato nei modelli OLS e con i *matching estimators* sia in parte sovrastimato a causa dei problemi di selezione.

Tabella 6. L'impatto delle ore di formazione sui crediti formativi maturati dagli studenti iscritti al Corso di Laurea in Economia Aziendale. Regressioni su diversi campioni di studenti

	(1)	(2)	(3)	(4)
	Tutti gli Studenti	Campione Test	Campione Test + Questionario	Sondaggio Studenti
Partecipazione80%	8.469*** (1.513)	7.847*** (1.614)	6.821*** (1.831)	3.167* (1.912)
Donna	0.619 (1.457)	1.650 (1.491)	0.913 (1.722)	-2.175 (1.537)
Voto Maturità	0.730*** (0.081)	0.620*** (0.090)	0.685*** (0.104)	0.814*** (0.103)
Scientifico	17.125*** (2.044)	13.752*** (2.217)	15.982*** (2.560)	11.416*** (2.659)
Classico	7.107* (3.795)	8.037* (4.392)	8.926* (4.852)	12.468*** (4.577)
Istituto Tecnico Commerciale	8.730*** (1.988)	8.139*** (2.043)	8.555*** (2.271)	6.599*** (2.427)
Liceo Socio-Psico-Ped	0.821 (4.646)	0.829 (4.462)	2.320 (5.383)	0.425 (4.548)
Anno diploma	-1.302 (1.253)	-1.459 (1.359)	-0.676 (1.291)	0.510 (1.771)
Età	-2.097* (1.244)	-2.339* (1.324)	-1.573 (1.242)	-2.128* (1.287)
Percentuale Risposte Corrette Matematica		22.380*** (4.437)	19.347*** (5.045)	25.756*** (5.037)
Percentuale Risposte Corrette Italiano		-0.157 (4.610)	-1.721 (5.086)	-1.852 (4.781)
Istruzione Padre			-0.004 (0.235)	
Istruzione Madre			-0.075 (0.239)	
Costante	2,609.532 (2,542.251)	2,930.117 (2,754.038)	1,339.928 (2,615.595)	-1,037.067 (3,573.234)
Osservazioni	462	441	370	306
R-quadro	0.374	0.410	0.390	0.413

Stime OLS sugli studenti di Economia Aziendale. Variabile dipendente: Crediti maturati dagli studenti. In tutte le regressioni controlliamo per dummies di provincia di residenza, età e anno di conseguimento della maturità. Gli *standard error* sono robusti all'eteroschedasticità. I simboli ***, **, * indicano, rispettivamente, significatività al 1%, 5% e 10%.

7. Considerazioni conclusive

In questo lavoro abbiamo cercato di comprendere se è possibile intervenire con politiche inserite in uno stadio avanzato del processo formativo per cercare di accrescere le competenze di base degli individui e per facilitare l'ulteriore acquisizione di capitale umano. A tale scopo abbiamo considerato un progetto finanziato dalla Regione Calabria che nell'anno accademico

2008-09 ha offerto agli studenti immatricolati all'Università della Calabria una serie di attività formative volte a colmare eventuali lacune nelle competenze di base.

Allo scopo di analizzare l'effetto prodotto da questo intervento sono state considerate alcune misure di successo accademico, quali il numero di crediti maturati e il voto medio ottenuto agli esami nel corso del primo anno di Università. I risultati delle stime OLS basate su tutti gli studenti immatricolati nell'anno accademico 2008-09 all'Università della Calabria mostrano che la partecipazione all'intervento è associata a una migliore performance accademica sia in termini di conseguimento di crediti, sia in termini di voto medio agli esami. Le stime ottenute usando tecniche di *matching* confermano sostanzialmente i risultati ottenuti con il modello di regressione.

Nonostante nelle stime si è controllato per un ampio numero di caratteristiche osservabili, non è possibile escludere che l'effetto stimato sia dovuto a fattori confondenti che non permettono di individuare un rapporto di causalità. Gli studenti che hanno partecipato all'intervento potrebbero essere caratterizzati da particolari caratteristiche inosservabili (ad esempio, una più forte motivazione) che sono correlate anche con una migliore performance. Per cercare di ovviare a questi problemi abbiamo effettuato la nostra analisi anche su campioni di studenti più omogenei che potrebbero avere livelli di motivazione simili. In particolare, abbiamo seguito l'approccio suggerito da Leuven e Oosterbeeck (2008) e abbiamo considerato come gruppo di controllo solo gli studenti che non hanno partecipato al progetto per motivi casuali. L'impatto stimato si riduce, ma è ancora rilevante: gli studenti che hanno seguito almeno l'80% delle attività formative previste nell'ambito dell'intervento di policy conseguono circa 3 crediti in più rispetto agli studenti che non hanno seguito i corsi a causa di fattori puramente casuali. Pertanto, sembrerebbe che le attività formative del progetto hanno avuto un impatto positivo, ma i problemi di selezione tendono a generare una sovrastima degli effetti.

Un interessante oggetto di analisi, da sviluppare in ricerche future, è una valutazione costi-benefici dell'intervento, cioè un'analisi che miri a misurare se i benefici (diretti e indiretti) derivanti dal progetto siano superiori ai suoi costi.

Riferimenti bibliografici

- Abadie, A., Drukker, D., Leber Herr, J., e Imbens, G., (2004), "Implementing Matching Estimators for Average Treatment Effects in Stata", *Stata Journal*, 4, pp. 290–311.
- Abadie, A., e Imbens, G. (2006), "Large Sample Properties of Matching Estimators for Average Treatment Effects", *Econometrica*, 74, pp. 235–67.
- Angrist J., Lavy V. (1999), "Using Maimonides' rule to estimate the effect of class size on scholastic achievement", *Quarterly Journal of Economics*, 114 (2), pp. 533-575.
- Angrist, J. (1998), "Estimating The Labor Market Impact of Voluntary Military Service Using Social Security Data on Military Applicants", *Econometrica*, 66(2), pp. 249-288.
- Ashenfelter, O., (1978), "Estimating the Effects of Training Programs on Earnings," *Review of Economics and Statistics*, 60:1 (February), 47–57.
- Ashenfelter, O., and D. Card, (1985), "Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs," *Review of Economics and Statistics*, 67:4, 648–660.
- Becker S., Ichino, A. (2002), "Estimation of Average Treatment Effects Based on Propensity Scores", *The Stata Journal*, Vol.2, No.4, pp. 358-377.
- Bettinger E. and Calcagno J. (2009) The Impact of Postsecondary Remediation Using a Regression Discontinuity Approach: Addressing Endogenous Sorting and Noncompliance, mimeo.
- Bettinger E. and Long B., (2009), Addressing the needs of underprepared students in higher education: does college remediation work?, *Journal of human resources*, 44(3), pp.736-771.
- Card D., Krueger A. (1996). "School Resources and Student Outcomes: An Overview of the Literature and New Evidence from North and South Caroline", *Journal of Economic Perspectives*, 10, 4, pp. 31-50.
- Carneiro, P., Heckman, J., (2003), "Human Capital Policy", IZA Discussion Paper No. 821.
- Dehejia, R., Wahba, S. (1999), "Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs", *Journal of the American Statistical Association*, 94, 448, pp. 1053-1062.
- Duflo E., Glennerster R. and Kremer M. (2007), "Using Randomization in Development Economics Research: A Toolkit", CERP Discussion Paper n. 6059.
- Fuchs, T., Wöessmann, L., (2007), "What accounts for international differences in student performance? A re-examination using PISA data", *Empirical Economics*, 32:433–464
- Glewwe, P. and Kremer M. (2006) "Schools, teachers, and educational outcomes in developing countries", In *Handbook of the Economics of Education*, edited by Eric A. Hanushek and Finis Welch. Amsterdam: North Holland:943-1017.
- Hanushek E. (1986), "The economics of schooling: Production and efficiency in public schools", *Journal of Economic Literature*, 24, pp. 1141–1177.
- Hanushek, E. (1999), What education production functions really show: a positive theory of education expenditures, *Economics of Education Review*, 18(2), pp. 223-239.
- LaLonde, R. (1986), "Evaluating the Econometric Evaluations of Training Programs," *American Economic Review* 76:4, 604–620.
- Leuven, E. e Oosterbeek, H., (2008), "An Alternative Approach to Estimate the Wage Returns to Private-Sector Training", *Journal of Applied Econometrics*, 23: 423–434.
- Levin H. and Calcagno J. (2008), Remediation in the Community College: An Evaluator's Perspective, mimeo
- Rosenbaum P.R., Rubin D. B. (1983), "The central role of the propensity score in observational studies for causal effects", *Biometrika*, 70, pp. 41–45.
- Rubin D. (1977), "Assignment to Treatment Group on the Basis of a Covariate", *Journal of Educational Statistics*, 2(1), pp. 1-26.
- Stinebrickner, R. Stinebrickner, T. (2008), "The Causal Effect of Studying on Academic Performance," *The B.E. Journal of Economic Analysis & Policy*: Vol. 8: Iss. 1 (Frontiers).

- Todd P, Wolpin K. (2003), "On the specification and estimation of the production function for cognitive achievement", *Economic Journal*, 113, 3–33.
- Webbink, D., (2005), "Causal Effects in Education", *Journal of Economic Surveys*, 19, 4, pp. 535-560.